

1/5/5 (Item 5 from file: 351)
DIALOG(R) File 351:Derwent WPI
(c) 2006 The Thomson Corp. All rts. reserv.

013876961 **Image available**
WPI Acc No: 2001-361173/ 200138
XRPX Acc No: N01-262877

Information processor for vocal interaction translation system, has
speech recognition section which changes specific words in speech
recognition result, based on indication from machine translation section

Patent Assignee: SONY CORP (SONY)
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 2001100784	A	20010413	JP 99277744	A	19990930	200138 B

Priority Applications (No Type Date): JP 99277744 A 19990930

Patent Details:
Patent No Kind Lan Pg Main IPC Filing Notes
JP 2001100784 A 12 G10L-015/18

Abstract (Basic): JP 2001100784 A

NOVELTY - A machine translation section (2) indicates a speech
recognition section (1) for charging specific control in the
speed-recognition result, when machine translation of specific words in
the speech recognition result is not able to be performed. The speech
recognition section then charges the specific words in speech
recognition result, based on indication and then outputs to machine
translation section.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the
following:

- (a) Information processing procedure;
- (b) Recording medium

USE - For vocal interaction translation system. For
natural-language processing.

ADVANTAGE - Enables easily performing accurate natural language
processing operation.

DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of
vocal interaction translation system. (Drawing includes non-English
language text).

Speech recognition section (1)
Machine translation section (2)
pp; 12 DwgNo 1/9

Title Terms: INFORMATION; PROCESSOR; VOICE; INTERACT; TRANSLATION; SYSTEM;
SPEECH; RECOGNISE; SECTION; CHANGE; SPECIFIC; WORD; SPEECH; RECOGNISE;
RESULT; BASED; INDICATE; MACHINE; TRANSLATION; SECTION

Derwent Class: P86; T01; W04

International Patent Class (Main): G10L-015/18

International Patent Class (Additional): G06F-017/28; G10L-015/00;

G10L-015/22; G10L-015/28

File Segment: EPI; EngPI

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-100784

(43)Date of publication of application : 13.04.2001

(51)Int.Cl.

G10L 15/18
G06F 17/28
G10L 15/00
G10L 15/28
G10L 15/22

(21)Application number : 11-277744

(22)Date of filing : 30.09.1999

(71)Applicant : SONY CORP

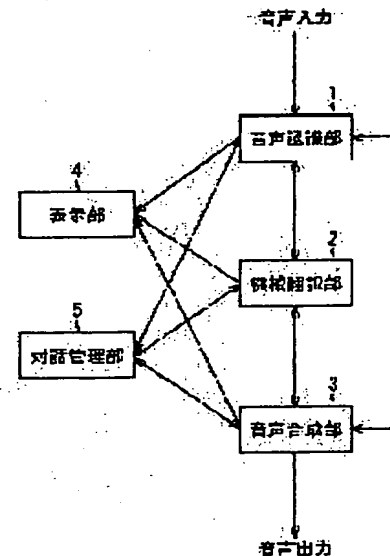
(72)Inventor : OMOTE MASANORI
OGAWA HIROAKI
HONDA HITOSHI
TSUTSUMI HIRONAGA

(54) INFORMATION PROCESSOR, INFORMATION PROCESSING METHOD AND RECORDING MEDIUM

(57)Abstract

PROBLEM TO BE SOLVED: To easily perform natural language processing high in precision.

SOLUTION: A speech recognition part 1 recognizes an inputted speech and supplies the speech recognition result to a machine translation part 2. The machine translation part 2 performs machine translation of the speech recognition result from the speech recognition part 1. If the speech recognition result is not constituted in the unit of components proper for machine translation, the machine translation part 2 instructs the speech recognition part 1 to change the unit of components. On the basis of the instruction from the machine translation part 2, the speech recognition part 1 changes the unit of components of the speech recognition result, for example, from the unit of clauses to the unit of words to output the result to the machine translation part 2.



音声対話/翻訳システム

LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application converted
registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection]

[Date of requesting appeal against examiner's decision of
rejection]

[Date of extinction of right]

SPR
IPR

(10) 日本国特許庁 (F) (12) 公開特許公報 (A)

(11) 特許公開番号
特開2001-100784
(P2001-100784A)
(12) 公開日 平成13年4月13日 (2001.4.13)

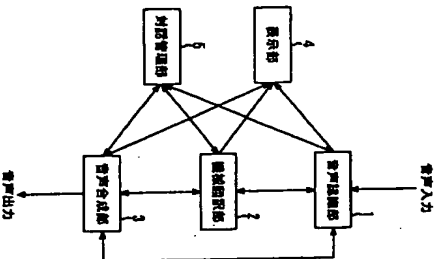
(6) In.CI.	識別記号	F I	チヨビ(参考)
G10L 15/18		G10L 3/00	637J 5B091
G06F 17/28		G06F 16/38	V 6D016
G10L 15/00		G10L 3/00	651C
15/28			681H
16/22			671T

審査請求 未請求 請求項の数15 OL (全12頁)

(21) 出願番号	特願平11-277744	(71) 出願人	00002185 ソニー株式会社 東京都品川区北品川6丁目7番5号 ソニー株式会社内
(22) 出願日	平成11年9月30日 (1999.9.30)	(72) 発明者	森 雅則 東京都品川区北品川6丁目7番5号 ソニー株式会社内
		(73) 発明者	小川 浩明 東京都品川区北品川6丁目7番5号 ソニー株式会社内
		(74) 代理人	100082131 弁理士 橋本 健雄

最終頁に続く

- (54) 【発明の名称】 情報処理装置および情報処理方法、並びに記録媒体
- (67) 【要約】
- 【課題】 容易に、精度の高い自然言語処理を行う。
- 【解決手段】 音声認識部1では、そこに入力される音声から音声認識結果、機械翻訳部2では、音声認識部1からの音声認識結果が機械翻訳される。この場合において、機械翻訳部2は、音声認識結果が、機械翻訳を行うのに適切な構成単位で構成されていない場合、その構成単位を変更するように、音声認識部1に指示する。音声認識結果の構成単位を、例えば、文節単位から単語単位に変更し、機械翻訳部2に出力する。



音声認識/翻訳システム

1 (2) 特開2001-100784

【特許請求の範囲】

【請求項1】 自然言語処理を行う自然言語処理装置に
対して、自然言語処理の対象とする対象データを出力す
る情報処理装置であって、
入力データを処理して得られる、所定の構成単位で構成
される前記対象データを、前記自然言語処理装置に出力
する出力手段と、
前記自然言語処理装置からの指示に応じて、前記対象デ
ータを構成する構成単位を変更する変更手段とを含むこ
とを特徴とする情報処理装置。

【請求項2】 前記入力データは音声データであり、
前記対象データは、前記音声データに基づいて行われた
音声認識結果であることを特徴とする請求項1に記載の
情報処理装置。

【請求項3】 前記音声データに基づいて音声認識を行
い、その音声認識結果を、前記対象データとして出力す
る音声認識手段をさらに含むことを特徴とする請求項2
に記載の情報処理装置。

【請求項4】 前記音声認識手段は、前記音声認識結果
を構成する構成単位を変更するための変更情報を含む所
書または文法規則を用いて音声認識を行い、
前記変更手段は、前記変更情報に基づいて、前記音声認
識結果を構成する構成単位を変更することを特徴とする
請求項3に記載の情報処理装置。

【請求項5】 前記自然言語処理装置は、前記音声認識
結果に対する返答を生成して対話を行うための管理を行
う対話管理装置、または前記音声認識結果を機械翻訳す
る機械翻訳装置であることを特徴とする請求項1に記載
の情報処理装置。

【請求項6】 前記入力データを自然言語処理し、その
自然言語処理結果を、前記対象データとして出力する自
然言語処理手段をさらに含むことを特徴とする請求項1
に記載の情報処理装置。

【請求項7】 前記自然言語処理手段は、前記入力デー
タに対する返答を生成して対話を行うための対話管理手
段、または前記入力データを機械翻訳する機械翻訳装置
を行うことを特徴とする請求項6に記載の情報処理装
置。

【請求項8】 前記自然言語処理装置をさらに含むこと
を特徴とする請求項1に記載の情報処理装置。

【請求項9】 自然言語処理を行う自然言語処理装置に
対して、自然言語処理の対象とする対象データを出力す
る情報処理方法であって、
入力データを処理して得られる、所定の構成単位で構成
される前記対象データを、前記自然言語処理装置に出力
する出力手段と、
前記自然言語処理装置からの指示に応じて、前記対象デ
ータを構成する構成単位を変更する変更手段とを含むこ
とを特徴とする情報処理方法。

【請求項10】 自然言語処理を行う自然言語処理装置

に対して、自然言語処理の対象とする対象データを出力
する処理を、コンピュータに行わせるプログラムが記録
されている記録媒体であって、
入力データを処理して得られる、所定の構成単位で構成
される前記対象データを、前記自然言語処理装置に出力
する出力手段と、
前記自然言語処理装置からの指示に応じて、前記対象デ
ータを構成する構成単位を変更する変更手段とを含むこ
とを特徴とする記録媒体。

【請求項11】 自然言語処理の対象とする対象データ
を出力する出力装置からの前記対象データを受信し、自
然言語処理を行う情報処理装置であって、
前記対象データを構成する構成単位が、自然言語処理が
可能な単位であるかどうかを判定する判定手段と、
前記判定手段の判定結果に基づいて、前記対象データを
構成する構成単位の変更を、前記出力装置に対して指示
する指示手段とを含むことを特徴とする情報処理装置。

【請求項12】 前記出力装置は、音声認識装置または
自然言語処理を行い、その処理結果を、前記対象データ
として出力することを特徴とする請求項11に記載の情報
処理装置。

【請求項13】 前記出力装置をさらに含むことを特徴
とする請求項11に記載の情報処理装置。

【請求項14】 自然言語処理の対象とする対象データ
を出力する出力装置からの前記対象データを受信し、自
然言語処理を行う情報処理方法であって、
前記対象データを構成する構成単位が、自然言語処理が
可能な単位であるかどうかを判定する判定手段と、
前記判定手段の判定結果に基づいて、前記対象データを
構成する構成単位の変更を、前記出力装置に対して指示
する指示手段とを含むことを特徴とする情報処理方法。

【請求項15】 自然言語処理の対象とする対象データ
を出力する出力装置からの前記対象データを対象とした
自然言語処理を、コンピュータに行わせるプログラムが
記録されている記録媒体であって、
前記対象データを構成する構成単位が、自然言語処理が
可能な単位であるかどうかを判定する判定手段と、
前記判定手段の判定結果に基づいて、前記対象データを
構成する構成単位の変更を、前記出力装置に対して指示
する指示手段とを含むことを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、情報処理装置およ
び情報処理方法、並びに記録媒体に関し、特に、例え
ば、音声認識装置の出力を、自然言語処理装置に処理さ
せる場合等において、音声認識装置の出力の構成単位
を、自然言語処理装置の処理に適応的な形にするように変

更すること、精度の高い処理を行うことができるようにする情報処理装置および情報処理方法、並びに記録媒体に関する。

【0002】従来の技術 例えば、音声による対話を行う音声対話システムや、ある音源による音声、他の音源による音声に同期する音声認識システムでは、入力された音声、音声認識する音声認識装置を必要の構成要素として、その音声認識装置による音声認識結果が自然言語処理される。

【0003】即ち、音声対話システムでは、音声認識装置による音声認識結果に対する返答を生成して対話を行うための管理を行う対話管理装置によって、対話管理が行われる。また、音声認識システムでは、音声認識装置による音声認識結果を機械翻訳する機械翻訳装置によって機械翻訳処理が行われる。

【0004】【発明が解決しようとする課題】ところで、音声対話システムや音声認識システムが高度化、複雑化すると、音声認識装置と、対話管理装置や機械翻訳装置とは、別々に関与されることが多くなる。この場合、音声認識装置は、音声、精度良く認識することができるとするようにならなければならない。一方、対話管理装置や機械翻訳装置は、対話管理処理や機械翻訳処理を精度良く行うことができるように入力力が与えられることを前提として、そのような入力の処理を行うように構成される。

【0005】具体的には、例えば、連続音声や精度良く認識するために、幾つかの単語をまとめた単位で処理するのが好ましく、従って、音声認識装置では、そのように幾つかの単語をまとめた単位の音声認識結果が出力される。一方、機械翻訳を精度良く行うためには、入力力が形態素単位で与えられるのが好ましく、従って、機械翻訳装置は、入力力が形態素単位で与えられるものとして構成される。

【0006】従って、上述のような音声認識装置と、対話管理装置や機械翻訳装置等の自然言語処理装置とを、部品として用いて、音声対話システムや音声認識システムを構成した場合、自然言語処理装置では、音声認識装置が出力する音声認識結果を解析し、自然言語処理可能な単位（精度の良い自然言語処理を行うのに適切な単位）に変更する必要がある。処理負担が大となる。

【0007】本発明は、このような状況に鑑み、なされたものであり、容易に、精度の高い自然言語処理を行うことができるようにするものである。

【0008】【課題を解決するための手段】本発明の第1の情報処理装置は、入力データを処理して得られる、所定の構成単位で構成される対象データを、自然言語処理装置に出力する出力手段と、自然言語処理装置からの指示に応じ

60

て、対象データを構成する構成単位を変更する変更手段とを含むことを特徴とする。

【0009】入力データは音声データとすることができ、対象データは、音声データに基づいて行われた音声認識結果とすることができ、

【0010】第1の情報処理装置には、音声データに基づいて音声認識を行い、その音声認識結果を、対象データとして出力する音声認識手段をさらに設けることができる。

【0011】音声認識手段には、音声認識結果を構成する構成単位を変更するための変更情報を含む情報または文法規則を用いて音声認識を行わせ、変更手段には、変更情報に基づいて、音声認識結果を構成する構成単位を変更させることができる。

【0012】自然言語処理装置は、音声認識結果に対する返答を生成して対話を行うための管理を行う対話管理装置、または音声認識結果を機械翻訳する機械翻訳装置とすることができ、

【0013】また、第1の情報処理装置には、入力データを自然言語処理し、その自然言語処理結果を、対象データとして出力する自然言語処理手段をさらに設けることができる。

【0014】自然言語処理手段には、入力データに対する返答を生成して対話を行うための対話管理処理、または入力データを機械翻訳する機械翻訳処理を行わせることができる。

【0015】第1の情報処理装置には、自然言語処理装置をさらに設けることができる。

【0016】本発明の第1の情報処理方法は、入力データを処理して得られる、所定の構成単位で構成される対象データを、自然言語処理装置に出力する出力システムと、自然言語処理装置からの指示に応じて、対象データを構成する構成単位を変更する変更手段とを含むことを特徴とする。

【0017】本発明の第1の記録媒体は、入力データを処理して得られる、所定の構成単位で構成される対象データを、自然言語処理装置に出力する出力システムと、自然言語処理装置からの指示に応じて、対象データを構成する構成単位を変更する変更手段とを含むプログラムが記録されていることを特徴とする。

【0018】本発明の第2の情報処理装置は、対象データを構成する構成単位が、自然言語処理可能な単位であるかどうかを判定する判定手段と、判定手段の判定結果に基づいて、対象データを構成する構成単位の変更を、出力装置に対して指示する指示手段とを含むことを特徴とする。

【0019】出力装置には、音声認識処理または自然言語処理を行わせ、その処理結果を、対象データとして出力力とさせることができる。

【0020】また、第2の情報処理装置には、出力装置

をさらに設けることができる。

【0021】本発明の第2の情報処理方法は、対象データを構成する構成単位が、自然言語処理可能な単位であるかどうかを判定する判定ステップと、判定ステップにおける判定結果に基づいて、対象データを構成する構成単位の変更を、出力装置に対して指示する指示システムとを含むことを特徴とする。

【0022】本発明の第2の記録媒体は、対象データを構成する構成単位が、自然言語処理可能な単位であるかどうかを判定する判定ステップと、判定ステップにおける判定結果に基づいて、対象データを構成する構成単位の変更を、出力装置に対して指示する指示システムとを含むプログラムが記録されていることを特徴とする。

【0023】本発明の第1の情報処理装置および情報処理方法、並びに記録媒体においては、入力データを処理して得られる、所定の構成単位で構成される対象データが、自然言語処理装置に出力され、自然言語処理装置からの指示に応じて、対象データを構成する構成単位が変更される。

【0024】本発明の第2の情報処理装置および情報処理方法、並びに記録媒体においては、対象データを構成する構成単位が、自然言語処理可能な単位であるかどうかを判定され、その判定結果に基づいて、対象データを構成する構成単位の変更が、出力装置に対して指示される。

【0025】

【発明の実施の形態】図1は、本発明を適用した音声対話/翻訳システム（システム）とは、複数の装置が論理的に集合したものをいい、各装置の装置が同一筐体中にあるか否かは問わない）の一実施の形態の構成例を示している。

【0026】この音声対話/翻訳システムでは、音声が入力されると、その音声に対する返答が出力されたり、また、その音声の対話出力されるようになっている。さらに、入力された音声に対する返答を、その音声の音源以外の音源に翻訳して出力することもできるようになっている。

【0027】即ち、音声認識部1には、例えば日本語などによる音声が入力されるようになっており、音声認識部1は、入力された音声や音源を認識し、その音声認識結果としてのテキスト、その他付随する情報を、機械翻訳部2や、表示部4、対話管理装置5などに、必要に応じて出力する。

【0028】機械翻訳部2は、音声認識部1が出力する音声認識結果を解析し、入力された音声の音源以外の、例えば英語に機械翻訳し、その翻訳結果としてのテキスト、その他付随する情報を、音声合成部3や、表示部4、対話管理装置5などに、必要に応じて出力する。音声合成部3は、機械翻訳部2や対話管理装置5などの出力に基づいて音声合成処理を行い、これにより、入力された

音声に対する返答、あるいはその音声の、他の音源への翻訳結果としての合成音を出力する。

【0029】表示部4は、例えば、液晶ディスプレイ等で構成され、音声認識部1による音声認識結果や、機械翻訳部2による機械翻訳結果、対話管理装置5が生成する返答等を、必要に応じて表示する。

【0030】対話管理装置5は、音声認識部1の音声認識結果に対する返答を生成し、機械翻訳部2や、音声合成部3、表示部4、対話管理装置5に、必要に応じて出力する。さらに、対話管理装置5は、機械翻訳部2の機械翻訳結果に対する返答を生成し、音声合成部3や、表示部4に、必要に応じて出力する。

【0031】以上のように構成される音声対話/翻訳システムにおいて、入力された音声に対する返答を出力する場合には、まず、その入力された音声は、音声認識部1で音声認識され、対話管理装置5に出力される。対話管理装置5では、音声認識部1による音声認識結果に対する返答が生成され、音声合成部3に供給される。音声合成部3では、対話管理装置5からの返答に対応する合成音が生成されて出力される。

【0032】また、入力された音声の対話を出力する場合には、その入力された音声は、音声認識部1で音声認識され、機械翻訳部2に供給される。機械翻訳部2では、音声認識部1による音声認識結果が機械翻訳され、音声合成部3に供給される。音声合成部3では、機械翻訳部2からの翻訳結果に対応する合成音が生成されて出力される。

【0033】さらに、入力された音声に対する返答を、その音声の音源以外の音源に翻訳して出力する場合に、まず、その入力された音声は、音声認識部1で音声認識され、対話管理装置5に出力される。対話管理装置5では、音声認識部1による音声認識結果に対する返答が生成され、機械翻訳部2に供給される。機械翻訳部2では、対話管理装置5からの返答が機械翻訳され、音声合成部3に供給される。音声合成部3では、機械翻訳部2からの翻訳結果に対応する合成音が生成されて出力される。

【0034】なお、入力された音声に対する返答を、その音声の音源以外の音源に翻訳して出力する場合には、音声認識部1による音声認識結果を、機械翻訳部2で機械翻訳し、その翻訳結果に対する返答を、対話管理装置5で生成し、合成音で出力するようにすることも可能である。

【0035】次に、図2は、図1の音声認識部1の構成例を示している。

【0036】ユーザの発話は、マイク11に入力され、マイク11では、その発話が、電気信号としての音声信号に変換される。この音声信号は、AD(Analog/Digital)変換部12に供給される。AD変換部12では、マイク11からのアナログ信号である音声信号がサンプリング

60

ズ、量子化され、デジタル信号である音声データに変換される。この音声データは、伸張抽出部13に供給される。

【0037】伸張抽出部13は、AD変換部12からの音声データについて、適当なフレームごとに、例えば、スベクトルや、線形予測誤差、ケプストラム係数、線形スベクトル等の特徴パラメータを抽出し、メモリ14に供給する。メモリ14では、伸張抽出部13からの特徴パラメータが一時記憶される。

【0038】ワッチング部15は、メモリ14に記憶された特徴パラメータに基づき、音声モデルデータベース16、辞書データベース17、および文法データベース18を必要に応じて参照しながら、ワイク11に入力された音声（入力音）を認識する。

【0039】即ち、音声モデルデータベース16は、音声認識する音声の言語における個々の音素や音節などの音素的な特徴を表す音声モデルを記憶している。ここで、音声モデルとしては、例えば、HMM(Hidden Markov Model)などを用いることができる。辞書データベース17は、認識対象の各単語について、その発音に関する情報や記憶された単語辞書を記憶している。文法データベース18は、辞書データベース17の単語辞書に登録されている各単語が、どのように連動する（つながるか）を記述した文法規則を記憶している。ここで、文法規則としては、例えば、文脈自由文法(CFG)やHPSG(Head-driven Phrase Structure Grammar)（主辞導向句構造文法）、統計的な単語連動確率(N-gram)などに基づく規則を用いることができる。

【0040】ワッチング部15は、辞書データベース17の単語辞書と参照することにより、音声モデルデータベース16に記憶されている音声モデルを接続することにより、単語の音声モデル（単語辞書）を構成する。さらに、ワッチング部15は、抽つかの単語辞書を、文法データベース18に記憶された文法規則を参照することにより接続し、そのようにして接続された単語辞書を用いて、特徴パラメータに基づき、例えば、HMM法等によって、ワイク11に与えられた音声を認識する。

【0041】そして、ワッチング部15による音声認識結果は、例えば、テキスト等で出力される。

【0042】なお、ワッチング部15は、機械翻訳部2や対話管理部6から、渡送する変換信号を受信することによって、変換信号を受信した場合、音声認識結果を構成する構成単位を変更して出力するようになっている。

【0043】即ち、ワッチング部15は、例えば、入力音声「これは何ですか」に対して、音声認識結果「これは何ですか」が得られた場合において、変換信号を受信したとき、1つの構成単位「これは何ですか」である音声認識結果を、3つの構成単位「これは、」「何で」「か」である音声認識結果や、5つの構成単位

「これ」、「は」、「何」、「です」、「か」である音声認識結果に変更して出力し直すようになっている。

【0044】このような構成単位の変更は、最初に得られた音声認識結果「これは何ですか」を構成する単語や単語を切り替えることで行うこともできるし、ワッチング部15による音声認識の処理単位を変更することで行うこともできる。

【0045】ワッチング部15による音声認識の処理単位を変更することにより、その音声認識結果の構成単位を変更する場合には、その変更のための情報を、辞書データベース17の単語辞書や、文法データベース18の文法規則に記述しておくことを行うことが可能である。

【0046】即ち、例えば、単語辞書に、文節「これは」と、その文節を構成する単語（形態素）「これ」と及び「は」とを対応付けて記述しておく。この場合、ワッチング部15では、単語辞書を参照することにより、入力音声「これは」に対して、「これは」という1つの構成単位で構成される音声認識結果を得ることも可能となるし、「これ」と「は」との2つの構成単位で構成される音声認識結果を得ることも可能となる。

【0047】なお、上述の場合には、文節と、その文節を構成する単語とを対応付けておくようにしたが、文と、その文を構成する文節とを対応付けておくことも可能であるし、さらに、その文節を構成する単語も対応付けておくことも可能である。

【0048】一方、文法データベース18の文法規則によつて、ワッチング部15による音声認識の処理単位を変更する場合に、例えば、主格が、代名詞と助詞「は」とを接続して構成されることを、文法規則に記述しておく。この場合も、ワッチング部15では、文法規則を参照することにより、代名詞「これ」と助詞「は」とからなる主格を表す入力音声「これは」に対して、「これは」という1つの構成単位で構成される音声認識結果を得ることも可能となるし、「これ」と「は」との2つの構成単位で構成される音声認識結果を得ることも可能となる。

【0049】ここで、ワッチング部15による音声認識の処理単位を変更し、単語辞書または文法規則のいずれか一方に基づいて行う他、その両方に基づいて行うことも可能である。さらに、ワッチング部15による音声認識の処理単位を変更し、単語辞書を複数用意するとともに、各単語辞書に対応する文法規則も用意し、変換信号に基づいて、音声認識に用いる単語辞書と文法規則の組合せを選択するようにすることを行うことも可能である。

【0050】なお、ワッチング部15による音声認識の処理単位を変更することにより、音声認識結果の構成単位を変更する場合には、ワッチング部15では、メモリ14に記憶された特徴パラメータを用いて、再度処理が

行われる。

【0051】次に、図3は、図1の機械翻訳部2の構成例を示している。

【0052】テキスト解析部21には、音声認識部1が出力する音声認識結果としてのテキストや、対話管理部5が出力する返答としてのテキストが、機械翻訳の対象として入力されるようになっており、テキスト解析部21は、辞書データベース24や解析用文法データベース25を参照しながら、そのテキストを解析する。

【0053】即ち、辞書データベース24には、各単語の表記や、解析用文法を適用するために必要な品詞情報などが記述された単語辞書が記憶されている。また、解析用文法データベース25には、単語辞書に記述された各単語の情報に基づいて、単語連動に関する制約等が記述された解析用文法規則が記憶されている。そして、テキスト解析部21は、その単語辞書や解析用文法規則に基づいて、そこに入力されるテキスト（入力テキスト）を構成する単語や構文情報等の言語情報を抽出する。ここで、テキスト解析部21における解析方法としては、例えば、正規文法や、文脈自由文法、HPSG、統計的な単語連動確率を用いたものなどがある。

【0054】テキスト解析部21で得られた入力テキストの解析結果としての言語情報は、言語変換部22に供給される。言語変換部22は、言語変換データベース26を参照し、入力テキストの言語の言語情報を、翻訳結果の言語の言語情報に変換する。

【0055】即ち、言語変換データベース26には、入力言語（言語変換部22への入力の言語）の言語情報から、出力言語（言語変換部22からの出力の言語）の言語情報への変換ルール（ラングルール）や、入力言語と出力言語との対照用例およびその対照例と入力言語との間の類似度の計算に用いられるシソーラス等の、言語情報を記憶するための言語変換データベース26が記憶されている。そして、言語変換部22では、このような言語変換データベースに基づいて、入力テキストの言語の言語情報を、出力言語の言語情報に変換される。

【0056】言語変換部22で得られた出力言語の言語情報は、テキスト生成部23に供給され、テキスト生成部23は、辞書データベース27および生成用文法データベース28を参照することにより、出力言語の言語情報から、入力テキストを出力言語に翻訳したテキストを生成する。

【0057】即ち、辞書データベース27には、出力言語の文を生成するために必要な単語の品詞や活用形等の情報や記述された単語辞書が記憶されており、また、生成用文法データベース28には、出力言語の文を生成するために必要な単語の活用規則や語順の制約等の生成用文法規則が記憶されている。そして、テキスト生成部23は、これらの単語辞書および生成用文法規則に基づいて、

言語変換部22からの言語情報を、テキストに変換して出力する。

【0058】なお、テキスト解析部21は、入力テキストの構成単位が、その入力テキストの解析可能な単位（入力テキストの解析を行うのに適切な単位）であるかどうかを判定し、可能な単位であれば、上述のように、入力テキストの解析を行う。一方、入力テキストの構成単位が、その入力テキストの解析可能な単位でない場合には、テキスト解析部21は、入力テキストの構成単位の変更を示す変換信号を、音声認識部1に送信する。音声認識部1では、上述したように、変換信号に基づいて、入力テキストとなる音声認識結果の構成単位が変更され、その結果、テキスト解析部21には、構成単位が変更された音声認識結果が、入力テキストとして供給される。そして、テキスト解析部21では、再び、その入力テキストの構成単位が解析可能な単位であるかどうかを判定され、以下、同様の処理が繰り返される。

【0059】ここで、対話管理部6も、機械翻訳部2と同様に、音声認識部1による音声認識結果を対象とした自然言語処理の1つである対話管理処理を行うが、その際、音声認識部1に対して、必要に応じて変換信号を出力することができるようになっている。

【0060】次に、図4は、図1の音声合成部3の構成例を示している。

【0061】テキスト解析部31には、機械翻訳部2が出力する翻訳結果としてのテキストや、対話管理部6が出力する返答としてのテキストが、音声合成部3の対象として入力されるようになっており、テキスト解析部31は、辞書データベース34や解析用文法データベース35を参照しながら、そのテキストを解析する。

【0062】即ち、辞書データベース34には、各単語の品詞情報や、読み、アクセント等の情報が記憶された単語辞書が記憶されており、また、解析用文法データベース35には、辞書データベース34の単語辞書に記述された単語について、単語連動に関する制約等の解析用文法規則が記憶されている。そして、テキスト解析部31は、この単語辞書および解析用文法規則に基づいて、そこに入力されるテキストの形態素解析や構文解析等の解析を行い、後段の規則合成部32で行われる規則音声合成に必要な情報を抽出する。ここで、規則音声合成に必要な情報としては、例えば、ポーズの位置や、アクセントおよびイントネーションを制御するための情報その他の韻律情報や、各単語の発音等の言語情報などがあ

る。

【0063】テキスト解析部31で得られた情報は、規則合成部32に供給され、規則合成部32では、音素片データベース36を用いて、テキスト解析部31に入力されたテキストに対応する合成音の音声データ（ディジタルデータ）が生成される。

【0064】即ち、音楽データベータ36には、例えば、CV(Consonant, Vowel)や、VCV、CVC等の形で音楽データが記憶されており、規則合成部32は、テキスト解析部31からの情報に基づいて、必要な音楽データを選択し、さらに、ボース、アクセント、イントネーション等を適切に付加することで、テキスト解析部31に入力されたテキストに対応する合成音の音楽データを生成する。

【0065】この音楽データは、DA変換部33に供給され、そこで、アナログ信号としての音声信号に変換される。この音声信号は、図示せぬスピーカに供給され、これにより、テキスト解析部31に入力されたテキストに対応する合成音が出力される。

【0066】次に、図6のプロフローチャートを参照して、図1の音声対話/翻訳システムの動作について、さらに説明する。

【0067】音声認識部1に対して、音声が入力されると、音声認識部1では、ステップS1において、その音声が入力されると、その音声認識結果としてのテキストが、機械翻訳部2または対話管理部5に出力され、ステップS2に進む。

【0068】ステップS2では、機械翻訳部2において、音声認識部1からのテキストを機械翻訳する翻訳処理が行われ、または、対話管理部5において、音声認識部1からのテキストに対する返答を生成する対話管理部5が行われ、ステップS3に進み、音声認識部1において、機械翻訳部2または対話管理部5から変更信号を受信したかどうか判定される。

【0069】ステップS3において、機械翻訳部2または対話管理部5から変更信号を受信したと判定された場合は、ステップS1に戻り、音声認識部1は、その変更信号に基づき、音声認識結果の構成単位が変更されるように、再度、音声認識を行い、その音声認識結果を、機械翻訳部2または対話管理部5に出力する。そして、以下、同様の処理が繰り返される。

【0070】また、ステップS3において、機械翻訳部2または対話管理部5から変更信号を受信していないと判定された場合は、機械翻訳部2または対話管理部5は、ステップS2における翻訳処理または対話処理の結果得られるテキストを、音声合成部3に出力し、ステップS4に進む。

【0071】ステップS4では、音声合成部3において、機械翻訳部2または対話管理部5からのテキストに対応する合成音が生成されて出力され、処理を終了する。

【0072】次に、図6のプロフローチャートを参照して、音声認識部1の動作について、さらに説明する。

【0073】音声認識部1は、音声認識部1に音声が入力されると、音声認識部1において、その音声の音声認識結果の構成単位を設定し、ステップS12に進

む。ここで、新たな音声が入力された直後においては、ステップS11では、所定のデフォルトの構成単位が設定される。

【0074】ステップS12では、音声認識部1において、入力音声の音声認識が行われ、ステップS13に進み、直前のステップS11で設定された構成単位の音声認識結果が、機械翻訳部2または対話管理部5に出力される。そして、ステップS14に進み、機械翻訳部2または対話管理部5から、変更信号を受信したかどうか判定される。ステップS14において、変更信号を受信したと判定された場合は、ステップS11に戻り、音声認識結果の構成単位が、変更信号に基づいて小さくまたは大きく設定され、即ち、具体的には、例えば、文節単位から単語単位に設定され、または逆に単語単位から文節単位に設定され、ステップS12に進み、以下、同様の処理が繰り返される。従って、この場合、音声認識部1では、変更信号に基づき、構成単位が小さくまたは大きくされた音声認識結果が、その後に行われるステップS13で出力される。

【0075】一方、ステップS14において、変更信号を受信していないと判定された場合は、処理を終了する。

【0076】次に、図7のプロフローチャートを参照して、機械翻訳部2および対話管理部5の動作について、さらに説明する。

【0077】機械翻訳部2または対話管理部5は、音声認識部1から音声認識結果としてのテキストを受信すると、ステップS21において、そのテキストの構成単位を解析する。そして、ステップS22に進み、その構成単位が、機械翻訳部2または対話管理部5で処理を行うに適切な単位であるかどうか判定される。

【0078】ここで、音声認識部1からの音声認識結果の構成単位が、機械翻訳部2または対話管理部5で処理を行うに適切な単位であるかどうかは、例えば、音声認識結果を形態素解析することにより判定することができ、また、その判定は、例えば、機械翻訳部2または対話管理部5で処理を行うに適切な単位の文字列を記憶しておき、その文字列と、音声認識結果の構成単位とを比較すること等によって行うことも可能である。

【0079】ステップS22において、音声認識部1からの音声認識結果としてのテキストの構成単位が、機械翻訳部2または対話管理部5で処理を行うに適切な単位でないとは判定した場合、ステップS23に進み、構成単位が適切となるように、即ち、構成単位が大きければ小さくし、小さければ大きくするように指示する変更信号が、音声認識部1に出力される。そして、音声認識部1から、その変更信号にしたがって構成単位が変更された音声認識結果が供給されるのを持って、ステップS21に戻り、以下、同様の処理が繰り返される。

【0080】一方、ステップS22において、音声認識部1からの音声認識結果としてのテキストの構成単位

が、機械翻訳部2または対話管理部5で処理を行うのに適切な単位であると判定された場合は、ステップS24に進み、機械翻訳部2または対話管理部5において、音声認識結果が処理される。

【0081】即ち、機械翻訳部2では、音声認識結果の翻訳が行われ、対話管理部5では、音声認識結果の意味解析が行われ、さらに、その意味解析結果に基づいて、返答(対話文)が生成される。そして、機械翻訳部2または対話管理部5では、その処理結果が出力され、処理を終了する。

【0082】以上のように、音声認識部1において、その後で自然言語処理を行う機械翻訳部2または対話管理部5からの指示に基づいて、出力する音声認識結果の構成単位を、自然言語処理を行うのに適切な単位に変更して出力するようにしたので、機械翻訳部2または対話管理部5では、容易に、精度の高い自然言語処理を行うことが可能となる。

【0083】また、従来の場合には、音声認識結果の構成単位が、自然言語処理を行うのに適切な単位でない場合に、精度の高い自然言語処理を行うには、音声認識結果を対象とした各種の解析処理が煩雑であるか、あるいは不要なそのような解析処理が煩雑であるか、あるいは不要な結果を処理する必要があることが可能となる。

【0084】なお、上述の場合においては、音声認識部1において、その後で処理を行う機械翻訳部2または対話管理部5からの指示に基づき、出力する音声認識結果の構成単位を変更するようにしたが、その他、例えば、対話管理部5で生成した返答としてのテキストを、機械翻訳部2で翻訳する場合などは、対話管理部5において、その後で処理を行う機械翻訳部2からの指示に基づき、出力する返答の構成単位を変更するようにすることが可能である。

【0085】また、音声認識部1の音声認識結果に対する返答を、対話管理部5で生成し、その返答を、機械翻訳部2で翻訳する場合には、音声認識部1において、その後で処理を行う機械翻訳部2および対話管理部5のいずれからの指示にも基づいて、出力する音声認識結果の構成単位を変更するようにすることも可能である。

【0086】次に、上述した一連の処理は、ハードウェアにより行うこともでき、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアとしての音声対話/翻訳システムに組み込まれているコンピュータや、汎用のコンピュータ等にインストールされる。

【0087】そこで、図8を参照して、上述した一連の処理を実行するプログラムをコンピュータにインストールし、コンピュータによって実行可能な状態とするために用いられる、そのプログラムが記憶されている記憶装置

体について説明する。

【0088】プログラムは、図8(A)に示すように、コンピュータ101に内蔵されている記憶媒体としてのハードディスク102や半導体メモリ103に予め記録しておくことができる。

【0089】あるいはまた、プログラムは、図8(B)に示すように、フロッピーディスク111、CD-ROM(Compact Disc Read Only Memory)112、MO(Magneto optical)ディスク113、DVD(Digital Versatile Disc)114、磁気ディスク115、半導体メモリ116などの記憶媒体に、一時的あるいは永続的に格納(記録)しておくことができる。このような記憶媒体は、いわゆるパッケージソフトウェアとして提供することができる。

【0090】なお、プログラムは、上述したような記憶媒体からコンピュータにインストールする他、図8(C)に示すように、ダウンロードサイト121から、デジタル衛星放送用の人工衛星122を介して、コンピュータ101に無線で転送したり、LAN(Local Area Network)、インターネットといったネットワーク131を介して、コンピュータ123に有線で転送し、コンピュータ101にインストールすることができる。

【0091】ここで、本明細書において、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に拘らず、時系列に処理する必要はなく、並列的あるいは個別に実行される処理(例えば、並列処理あるいはオブジェクトによる処理)を含むものである。

【0092】また、プログラムは、1のコンピュータにより処理されるものであっても良いし、複数のコンピュータにより分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

【0093】次に、図9は、図8のコンピュータ101の構成例を示している。

【0094】コンピュータ101は、図9に示すように、CPU(Central Processing Unit)142を内蔵している。CPU142には、バス141を介して、入力ライン145が接続されており、CPU142は、入力ライン145を介して、ユーザによって、出力インタフェース146を介して、ユーザによって、キーボードやマウス等で構成される入力部147が操作されることにより指令が入力されると、それにしたがって、図8(A)の半導体メモリ103に対応するROM(Random Only Memory)143に格納されているプログラムを実行する。あるいは、また、CPU142は、ハードディスク102に格納されているプログラム、衛星122若しくはネットワーク131から転送され、通信路148で受信されてハードディスク102にインストールされたプログラム、またはドライブ149に装着されたフロッピーディスク111、CD-ROM112、MOディスク11

15

3. DVD114、若しくは磁気ディスク115から読み出されてハードディスク102にインストールされたプログラムを、RAM(Random Access Memory)14にロードして実行する。そして、CPU142は、その処理結果を、例えば、入出力インタフェース145を介して、LCD(Liquid Crystal Display)等で構成される表示部146に、必要に応じて出力する。

【0086】

【発明の効果】本発明の第1の情報処理装置および情報処理方法、並びに記録媒体によれば、入力データを処理して得られる、所定の構成単位で構成される対象データが、自然言語処理装置に出力され、自然言語処理装置からの指示に応じて、対象データを構成する構成単位が変更される。従って、自然言語処理装置において、容易に、精度の高い自然言語処理を行うことが可能となる。

【0086】本発明の第2の情報処理装置および情報処理方法、並びに記録媒体によれば、対象データを構成する構成単位が、自然言語処理が可能である単位であるかどうか判定され、その判定結果に基づいて、対象データを構成する構成単位の変更が、出力装置に対して指示され、従って、容易に、精度の高い自然言語処理を行うことが可能となる。

【図面の簡単な説明】

【図1】本発明を適用した音声対話/翻訳システムの一実施形態の構成例を示すブロック図である。

【図2】音声認識部1の構成例を示すブロック図である。

【図3】機械翻訳部2の構成例を示すブロック図である。

【図4】音声合成部3の構成例を示すブロック図である。

【図5】図1の音声対話/翻訳システムの動作を説明す

るためのフローチャートである。

【図6】音声認識部1の動作を説明するためのフローチャートである。

【図7】機械翻訳部2および対話管理部6の動作を説明するためのフローチャートである。

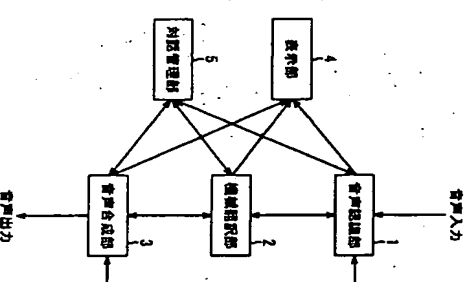
【図8】本発明を適用した記録媒体を説明するための図である。

【図9】図8のコМПE-101の構成例を示すブロック図である。

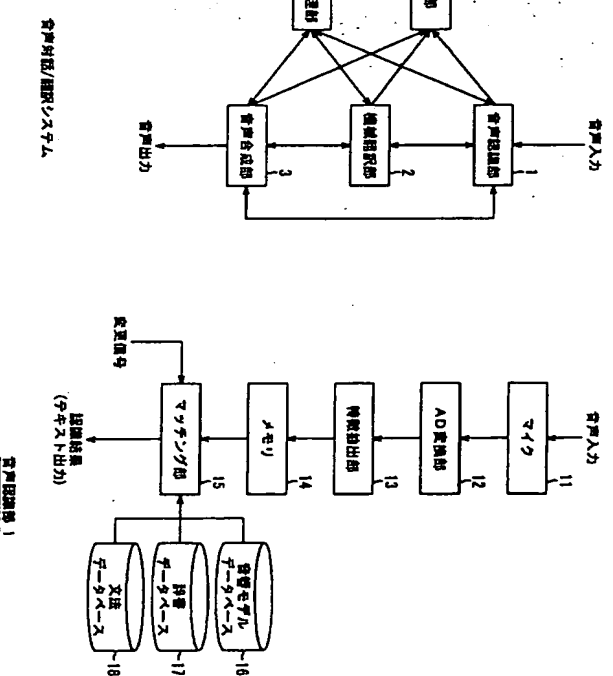
【符号の説明】

1 音声認識部、2 機械翻訳部、3 音声合成部、4 表示部、5 対話管理部、11 マイク(マイクログリップ)、12 AD変換部、13 特徴抽出部、14 メモリ、16 マシン学習部、16 音響モデルデータベース、17 辞書データベース、18 文法データベース、21 テキスト解析部、22 言語変換部、23 解析用文法データベース、24 辞書データベース、25 解析用文法データベース、26 言語変換データベース、27 辞書データベース、28 生成用文法データベース、31 テキスト解析部、32 規則合成部、33 DA変換部、34 辞書データベース、36 解析用文法データベース、102 ハードディスク、103 コンピュータ、102 ハードディスク、103 半導体メモリ、111 フロントディスプレイ、112 CD-ROM、113 MDディスク、114 DVD、115 磁気ディスク、116 半導体メモリ、121 タouchパネル、122 衛星、131 ネットワーク、141 バス、142 CPU、143 ROM、144 RAM、145 入出力インタフェース、146 表示部、147 入力部、148 通信部、149 ドライバ

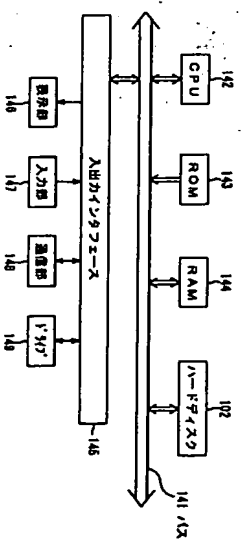
【図1】



【図2】



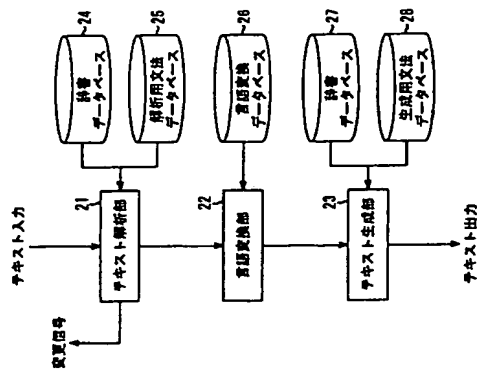
【図3】



【図9】

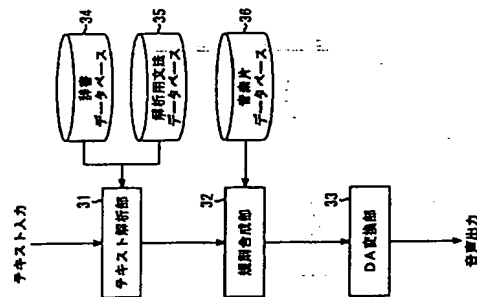
COMP-101

【図3】



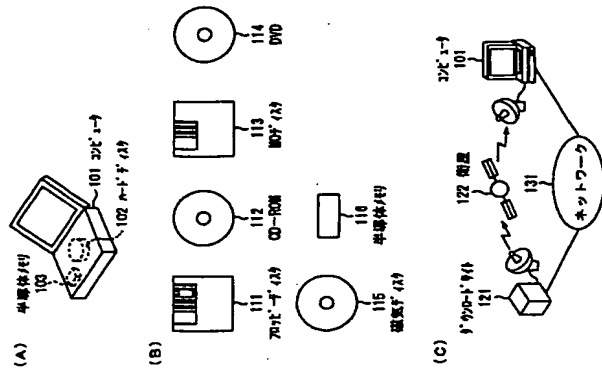
音声処理部 2

【図4】



音声処理部 3

【図8】



フロントページの続き

(72)発明者 本田 等
東京都品川区北品川6丁目7番35号 ソニ
株式会社内

(72)発明者 包 鉄良
東京都品川区北品川6丁目7番35号 ソニ
株式会社内

Fターム(参考) 58091 AA15 CA01 CB12 CB32

5D015 AA01 HH00 KK02 LL05 LL06